CONFIDENTIAL

SECRET

③

PHASE II FINAL REPORT

PROJECT
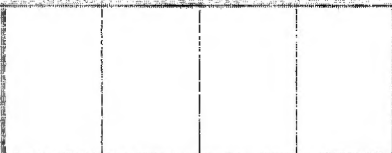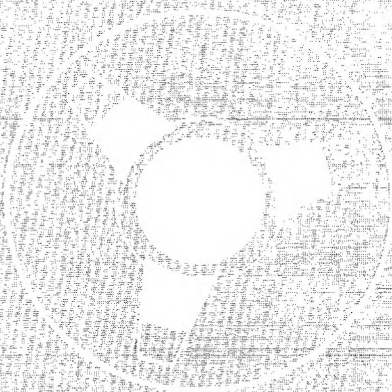# CHIVE

Volume III

IMPLEMENTATION PLAN

FOR INITIAL SYSTEM

CHIVE/R-3-65

1 March 1965

DIRECTORATE OF SCIENCE AND TECHNOLOGY

OFFICE OF COMPUTER SERVICES

SECRET

CONFIDENTIAL

U.S. Central Intelligence Agency,
Office of Computer Service.

Project CHIVE:

Phase II Final Report

Volume III

IMPLEMENTATION PLAN

FOR THE INITIAL SYSTEM

CHIVE/R-3-65

1 March 1965

TABLE OF CONTENTS

Page

SECRET CONFIDENTIAL

TABLES

CONFIDENTIAL

FIGURES

Chapter 3.1.

SUMMARY DESCRIPTION OF PHASE III PRODUCT

The basic goal of the Phase III effort is the creation
of a well-established and organized EDP environment
and the construction of an initial system component
which shall be judged ready for operational testing.
This component will be an experimental branch devoted

25X1A to work on ██████████████ The operational test period
should provide sufficient information to enable Agency
management to make a decision on whether or not to
proceed to full system implementation.

The initial operational component, so far as is possible
within the limitations of initial programming and
available manpower, should:

    a.   Provide a base leading toward the ultimate
          system objectives, i.e., broader coverage,
          more specific indexing, single point service,
          an all-source data base, capability to answer
          more complex questions, etc.

In addition, it should:

    b.   Have a limited, well-defined data base.

    c.   Benefit more than one customer group.

    d.   Provide the seed from which the total system
          can grow in an evolutionary, not revolutionary,
          manner.

                                SUMMARY DESCRIPTION
                                3.1.

CONFIDENTIAL

e.   Require a minimum of input and retrieval coor-
dination with the existing system.

f.   Offer a reasonably controlled environment for
testing both man/machine relations as well as
management techniques.

g.   Represent a wide range of the problems anticipated
in the final system.

h.   Not have a disrupting effect on current service
provided by OCR during the build-up period.

Indexing and file structuring concepts will be
tested and refined and the operational simulation period,
during the latter months of Phase III, should provide
sufficient information to assist in making a final
determination on the operational readiness of the initial
increment.

The initial design will not stress sophisticated or
unusual techniques per se. Humans will continue to do those
things they are able to do best. While the ultimate goal
may be extensive automation of the data input/output
process, the initial steps must settle for less--a
partially automated system utilizing a computer to
mechanize all tasks which clearly need not be performed
by hand.

While creation of a well-established and organized
EDP environment may require a heavy initial investment
with no immediate payoff when compared to the data base

SUMMARY DESCRIPTION
3.1.

-2-

SECRET

CONFIDENTIAL

to which the investment is applied, the long-range
objective of a broader data base, together with new
concepts in system information management, would
ultimately be satisfied with little additional investment
in EDP techniques.

3.1.1. RECOMMENDATIONS FOR INITIAL INCREMENT

A further description of the organization, manning
requirements, and tasks of the ███ Test Branch is
included in later sections of this volume. The discussion
which follows is included to set forth briefly the rationale
behind the decision to begin initial system activity
through the mechanism of the ███ element.

25X1A

25X1A

3.1.1.1. Geographic Approach

While CHIVE has argued, in earlier documentation, that
the central document/information reference system should
be organized along geographic lines, this view was not
automatically imposed on the initial system. Rather, a
number of possible implementation paths - by document,
by subject, by named object, and by area - were explored
before a final selection was made. These various alter-
native approaches are examined below.

SUMMARY DESCRIPTION
Recommendations
3.1.1.1.

3.1.1.2. Single Area Approach

In addressing the problem of selecting a point of
entry from which the full scale system could evolve,
the guidelines and objectives set forth at the beginning
of this volume were adopted.

The table below portrays in summary form the extent
to which various alternative initial system configurations
appeared to satisfy those guidelines and objectives.

Table 3-1

Alternative Initial Organization Configurations

Alternatives	Meets Objectives	Violates Objectives
1. All documents, all subjects, all countries	a, c, e, g	b, d, f, h
2. Single document type, all subjects, all countries	b, f, h *	a, c, d, e, g
3. All documents, single subject, all countries	g	a, b, c, d, e, f, h
4. All documents, single named object, all countries	d, f, g, h	a, b, c, e
5. All documents all subjects, single country	a, b, c, ** d, e f, g, h	

* Depends on document category chosen.

**Depends on country chosen.

SUMMARY DESCRIPTION
Recommendations
3.1.1.2.

Alternative 1 is obviously unrealistic since it literally requires implementation of the complete system right at the very start. This means we would be confronted by every type of documentary problem, all indexer support tools would have to have been developed in advance, all plans worked out for supplanting every component activity of the existing system, all personnel reassigned and retrained, and so forth. More specifically, such an approach would violate the objective of a limited, well-defined data base; it would require a revolutionary development program, it would fail to provide a suitable environment for testing purposes; and the effects of failure might well be calamitous.

Alternative 2, while it would offer the limited data base and controlled environment we desire, is unattractive because it would not incorporate some of our ultimate system objectives (e.g., all-source document handling), and it would serve only those customers interested in this one source. In essence, the SR entry point, for example, merely defers (but does not solve) the problem of how the system is to evolve. Once Comint materials were under control, we would be faced with the task of incrementing the system by source, which would be most difficult since

SUMMARY DESCRIPTION
Recommendations
3.1.1.2.

there is little to choose between the extremes of a
single document series (e.g., State airgrams) and all
collateral documents. Moreover, if we chose to evolve
by document categories, the problems of coordinating with
the existing system would be immense. Finally, the
system would be criticized for having failed to address
a sufficient variety of document problems which could
be anticipated in the full-scale system and, therefore,
having proved nothing about its long-run potential.

The notion represented by Alternative 3 is that
the system would process all documents reporting infor-
mation on all foreign countries, but that the subject
matter of these documents would have to pertain to a
single, though broad, subject area of intelligence interest,
e.g., economic affairs, military activities, or other.
Within the confines of this topical restriction, any
kind of data might be stored and retrieved, including

25X1B ████████████████████████ installations, locations,
activities, and events.

If we began with such a system, it would presumably
serve only one major component of the Agency, such as
OCI or OSI. Its most important disadvantage, however,
would be that the document base involved would inevitably

SUMMARY DESCRIPTION
Recommendations
3.1.1.2.

-6-

overlap with the existing system with consequent

duplication of activity. This we can ill afford in view

of the already tight manpower situation. Any request

on the new system, with the exception of those asking

only for the most recent information, would also have

to be levied on the old since the established files

could not be readily broken apart. In summary, the

complexity of operating a system of this kind concurrently

with the old would be great, its data base would be

larger than we would want to handle in the initial system,

and the consequences of failure would be serious.

Alternative 4 would give us a variety of documentary

inputs and the full spectrum of area problems. Only the

type of data to be processed and retrieved (e.g., ███████   25X1B

25X1B   ███████ organizational, or other) would be restricted.

Unfortunately, like Alternative 3, it would be impossible

to separate the documents on this basis, with the result

that both systems would receive essentially the same

inputs. Its other serious disadvantages would be that

the initial system could not test its integrated indexing

concept, the data base would be virtually that of the final

system, and it would require close coordination of request

handling with the established OCR systems.

SUMMARY DESCRIPTION
Recommendations
3.1.1.2.

Alternative 5 seems to offer the most advantages and fewest disadvantages of the options listed and was the one selected. Using this approach, a single country would be selected as the entry point but the entire range of documentary receipts on that country, and all types of subject matter would be handled. This would permit us to identify an initial geographical component of our total organization and test it in a real environment. The data base could be reasonably well defined, and if the country selected was of sufficiently widespread intelligence interest a good cross section of customers would be involved. The evolutionary path would be clear cut, involving the addition of new geographic area responsibilities to the new system. During the period of evolution, both the new and existing systems would have a clear understanding of their storage and retrieval obligations, and the need for inter-system communication and coordination would be minimized except in the case of "international" requests.

Certain problems stemming from linguistic variables and other processing problems peculiar to individual countries would obviously not be addressed in this kind of an initial system. However, none of the alternatives

SUMMARY DESCRIPTION
Recommendations
3.1.1 2.

-8-

proposed except that which implements the complete

system all at once would incorporate all the problems

which would be reflected in the final system.

3.1.1.3. Country Selection

For initial implementation purposes it was felt

that the country selected should have the following

characteristics:

- A manageably low document volume, probably
  somewhere between 50,000 and 100,000 items
  per year.

- Considerable intelligence significance to
  generate more than average consumer interest.

- Documentation which gives a sampling of major
  topic areas--i.e., political, scientific and
  technical, military, economic.

- Available personnel familiar with the geographic
  area and its documentation.

The USSR was eliminated from consideration immediately

since it would be dangerous to begin the operation with the

most strategic area of all, and because the volume of

receipts on the USSR would be extremely high. Most of

the West European, Near Eastern, Latin American and

African countries are either too weak in documentation

on major topics of general system interest, lack reporting

SUMMARY DESCRIPTION
Recommendations
3.1.1.3.

-9-

Next 4 Page(s) In Document Exempt

needed skills identified. The test branch would be brought
to operational strength around the tenth month of Phase III*
and will move into an experimental indexing and training
phase. (See timetable in Section 3.2.1.1.) Full system
testing and operational simulation would begin in month 19
and the ▓▓▓▓▓ branch would operate thereafter in parallel
with current OCR operations until a decision could be
made on the readiness of the initial component to assume
operational responsibilities.

To assist in the implementation of the test
component, and to provide logistical assistance and
operator advice to OCS, the CHIVE Support Staff (CSS)
should be enlarged during month 1 to a five-man team
consisting of four substantive analysts (probably
representing ▓▓▓▓▓▓▓▓▓▓▓▓▓ and one OCR support
programmer who should have a knowledge of MD or SR
maintained EAM authority files. Two additional OCR
support programmers should join the staff at a later
time. Personnel from GR and FDD could be committed on
an _ad hoc_ basis to work on inclusion of graphic indexes
and on the problem of obtaining select machine-readable
inputs from FDD and JPRS.

---

*Note: Phase III milestones are given in months numbered
serially from receipt of approval to proceed--month 1.

SUMMARY DESCRIPTION
Organization
3.1.2.

-15-

25X1A

25X1A

25X1B

25X1A

Both the CSS and ▦▦▦▦▦▦▦▦▦▦▦▦ could be slotted against OCR's T/O and would be under OCR management. The CSS, however, should be based in OCS/Development Division where work on the major system design problems and on development programming is assigned.

3.1.2.1. OCR Responsibilities

OCR should be responsible for guidance of the design activity, for establishing system boundaries and objectives and for review and acceptance of design decisions. The DD/I CHIVE Project Officer should be the responsible official for monitoring these functions. In addition, OCR should assume responsibility for providing the personnel of both the CSS and ▦▦▦ Test Branch, and for

25X1A

obtaining clearances for these people. Provision of

25X1A

space for the ▦▦▦ Branch as well as space in a secured area in which to establish the Document Delivery System, should be OCR's responsibilities. Personnel for the Document Delivery System could presumably be recruited from the present Library and SR document personnel and from MD's film personnel. Contracting and purchase of the Document System should also fall to OCR. Filming and key punch services related to implementation of the document system would also be OCR responsibilities.

SUMMARY DESCRIPTION
Organization
3.1.2.1.

3.1.2.2. OCS Responsibilities

Full design responsibility should rest with OCS. This includes system refinement and the execution and management of computer program design and development and system testing. In addition, OCS should be responsible for space to house the CHIVE design team, the CSS, the computer and attendant hardware. Contracting and costs of the computer equipment and optical character reader intended for system input should fall to OCS, as should contracting, housing, and management of the design and programming contract personnel. Contracts directly concerned with CHIVE should be coordinated with OCR. OCS should provide key punch and computer support to the system as it develops on a demand basis for requests and on a scheduled basis for input and maintenance.

3.1.2.3. Contractor Responsibilities

Under the guidance of OCS/DD, the contract personnel should be responsible for participating in and advising on all programming activities. This would include the development programming required for the initial system, and the programs required for extant OCR vocabulary control files designated for conversion to CHIVE formats.

SUMMARY DESCRIPTION
Organization
3.1.2.3.

The contractor would be expected to make recommendations
on program design, content and structuring, to propose
economical and feasible EAM file conversion routes, and to
continually bring to the system an awareness of new
programming or EDP developments which might benefit the
design activity. In addition, contract personnel should
participate in system review, and take part actively
in examination and evaluation of document delivery systems,
character readers and remote I/O devices, including
making recommendations on selection of equipment for
the system.

3.1.3. THE OPERATIONAL TESTING CONCEPT

The conclusion of the Phase III effort should
have modeled an initial CHIVE component which will be
ready for testing. The first testing stage, beginning
in month 19, will be a simulation of actual opera-
tional experience--a shakedown, in a sense, of the final
configuration derived from the experimentation of the
preceding months. Some further experimentation, adjust-
ment of programs, and refinement of procedures will
carry over into the initial testing stage. As far as
possible, however, the initial component will act like a
fully operational OCR component, i.e., it will receive,

SUMMARY DESCRIPTION
Testing Concept
3.1.3.

analyze, and index documents, forward index records to
the computer system, build files, and respond to queries.
The major difference between the ███ Branch and a   25X1A
"live" OCR component will be that it will not assume
actual operational responsibility to respond to queries,
since these will be borrowed, or captured at the time
of receipt, from the various OCR registers. It will
also use the old OCR files maintained by the registers.

The initial component must also prepare itself,
to contribute to such OCR summary files or scheduled
products as the ███████████████████████████████████
tribution only), ███████████████████████████████████   25X1B
In addition, the initial system must specify procedures
for preparation of finished intelligence reports and the
meeting of production deadlines for the scheduled products
cited above.

This testing will take place in parallel with
existing OCR operations for a span of six months or
longer, or until system and Agency management are satisfied
that the initial component is ready to accept opera-
tional responsibility. At that time ███ operations   25X1A
in the OCR divisions would be closed out and their per-

25X1B   sonnel and ███ and FIB-type working files transferred to

SUMMARY DESCRIPTION
Testing Concept
3.1.3.

25X1A

the new ▉▉▉ component. The manpower balance of the initial increment would not be increased, however, since the seasoned personnel would begin work on incrementation of the next geographic area as soon as the new personnel were trained.

Early experimentation, beginning after month 10, might involve an attempt to cope with the entire ▉▉▉▉ document flow on a "no-backlog" basis. (It is understood that backlogging might occur in an operation mode, but, as no query would be going on during indexing, an attempt to avoid backlogs should be made.) Plans would call for capturing a one month take of ▉▉▉ documents from all sources. Following the initial training period in concepts and procedures, the staff would be required to index these documents in one working month. If they were unable to do this, the indexing would be cut back and the identical documents would be indexed again within the identical time constraints. This process would be repeated until a match was obtained between document receipts and time available for indexing. At this point a second month's document take would be introduced and the entire process repeated to verify the successful indexing procedure.

25X1A

25X1A

SUMMARY DESCRIPTION
Testing Concept
3.1.3.

During this iterative indexing period, a basic set of query problems would be developed and passed against the indexed data following each iteration. Presumably a trade-off point would be reached between reduction of indexing and intelligibility of query response which would aid in the determination of the final indexing plan. Indexing levels for each iteration would be worked out in close cooperation with CSS, who would recommend exclusion standards and reflect requirements of their various OCR components, and with OCS/DD where index changes affected EDP requirements and specifications.

25X1A ████████Branch may assist in the development of such requirements and standards.

CHIVE's present organizational thinking postulates a group of geographic divisions staffed with area/topic-specialized information/query analysts who would mark elements of information in documents for later indexing by a large, separate staff of indexing personnel in a "Content Indexing Group" (see Vol. 5). Phase III should permit experimentation with this, and other, possible personnel configurations, but will not exclude the possibility of limiting its ultimate approach to document

25X1B control, saving the ████████/installation "analysis" problem for a later time.

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.

25X1B

25X1B

█████████ and organizational specialists from ██ and FIB could be brought into the test prior to the start of the general training period. These people would already have been given ISC training. They would go through the general CHIVE training program, but with particular emphasis on the query and file building aspects. Once experimental indexing was begun, however, their partici- pation would be directed toward specification of named object content for document indexing and on preparation

25X1B

of special ██████████ and installation information and summary files.

By the time the indexing/retrieval balance was tentatively established, the CTB would have reached full operational strength. The document flow would be started and routine indexing commence.

3.1.3.1.   Phasing in of New Areas

Since the planned system would increment by geographic

25X1A

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.1.

should have begun by mid-1968, with full integration of
the Far East completed by the end of 1969. At this point,
assuming that system concepts are working out as planned,
the major priority area, the USSR, could be attacked.
Eastern Europe might be the third point of incrementation,
followed by the Near East and Africa, Western Europe, and
Latin America.

As new geographic areas are added, a repetition of
many of the implementation tasks performed for the initial
component will be required, e.g., vocabulary control file
building (although not on the scale required for the
initial component), document and data selection, file
identification, and training.

Since CHIVE's recommended organizational configura-
tion assumes geographically oriented information/query
analysts marking elements of information in documents
for later coding by a large, possibly separate, indexing
staff, physical separation of the expanding indexing and
analytical staffs and their files may become a system
consideration as soon as a planned Far Eastern increment
is added to the initial component.

Here, the idea of OCR adopting a "pre-CHIVE"
configuration which will ease this transition becomes

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.1.

attractive. Following one possible line that OCR reorganiza-
tion might take, the CHIVE indexer staff, as it expands, could
become a component of an OCR "document processing center";
the information analyst staff could be based in an
information center--a component formed by merging present
OCR information functions under a single manager. As the
system grows, the personnel from the older document and
information components could move into the new activities
often without physical or managerial relocation. Trained
by co-workers, they would simply move from one processing
mode to another.

3.1.3.2. File Co-location

25X1A   No co-location of ▮▮▮▮ files with the initial
increment is planned for the operational test stage.
25X1B   Access to major document, ▮▮▮▮▮ and EAM files will
become a major problem only with the beginning of real
25X1A   operational activity by the ▮▮▮ component. OCR document
and microimage files could be placed under a single manage-
ment, and every effort made to centralize their storage
and maintenance. Failing this, repeated file relocation
will be required as new areas are added to the developing
system, and a large amount of analyst activity will be
required to tap all of the older, unrelocatable files

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.2.

which may be involved in query activity. This matter is elaborated somewhat in Section 3.2.3.1.

3.1.3.3. Exclusions

As indicated in Volume IV and in previous CHIVE documentation, a number of OCR functions are excluded from initial system attention. These include dissemination, acquisition, book cataloging and circulation, translation services, EAM support to other offices of the Agency, and the activities of OCR's liaison and administrative staffs.

Each of these functions presents a separate set of problems which will, in all likelihood, require separate solutions. They have been excluded from initial consideration because they are not central to the initial implementation plan. While OCS resources are available to OCR to attack any of these problems, concentration on these functions by the CHIVE Design Staff at this time could not but impede the course of initial system development.

One area which is of concern to system development will be explored during Phase III. This is the matter of translation services. CHIVE is in close contact with EDP developments on this front through the ALP project.

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.3.

However, the problem of obtaining selected translated materials as machine-readable system inputs is of significance as it impacts on the manpower requirements of the developing system, and a plan for obtaining such inputs should be ready for implementation prior to month 19. Some further discussion of how this activity might be carried out is available in Volume V.

Dissemination would probably be the next area to be addressed, and would involve the development of an automatic dissemination system, based on machine-stored user interest profiles which would be kept constantly under update and review. Automatic document dissemination could probably be linked to the inital header data indexing activity so that centralized dispatch of documents to CHIVE components, to user analysts, and to the central document store could be effected in one set of actions.

3.1.3.4. Inadequacies

The initial system, while it will model the total system objectives, will suffer from certain constraints imposed by development costs and rates and because of the existence of the old system. A heavy reliance on old document and EAM files can be anticipated through the first several years of operation, and this will require

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.4.

considerable manual labor in preparing products for system
customers. The variety of output products involved in
searches of the inherited or old files and new CHIVE files
will also prevent the production of a truly integrated
system response to many requests. Also, a variety of
query routes will have to be a part of the new system in
order to provide access to older EAM-based files. Specialists
in old file maintenance and query will have to be involved
in query activities on many occasions.

Limitations in the EDP environment of the new system
will also create certain hardships. Since a general
purpose program for query and file structuring will not be
available by the time of initial implementation, files
will be structured on the basis of requirements known
to exist in the various OCR registers. When unanticipated
types of system outputs are required, special programs will
have to be created by the information analyst specifying
his requirements to programmers or "EDP file analysts."
The need for special programming actions, however, need not
mean that special input processing will be required.

3.1.3.5. Initial EDP System Products

The initial EDP system will provide a capability to
produce an entire document index record--header data plus

<div align="right">
SUMMARY DESCRIPTION
Testing Concepts
3.1.3.5.
</div>

information indexed to the document--or to select combina-
tions of data from the index record, suppressing unwanted
indexing and arranging the required information in
meaningful combinations at the direction of the system
user. The initial system will provide machine assistance
in the production of extensive, updated, and non-redundant
formatted summary files of information on people, install-
ations and commodities or topics. One great power in the
EDP system is its ability to collect and arrange independent
data elements from many index records to aid the analyst
in producing non-redundant synthesized reports which can
be stored, updated or printed at the operator's command.
The system will also provide a limited capacity for
remote visual review of index records, will output
printed indexes, listings and summary reports, and will
retrieve documents and make them available in either
hard copy or microform.

3.1.3.6. Problems

The path of initial implementation, extending
through the Phase IV live operational tests, is beset
with a number of major problems. It is possible at
this time to suggest tentative solutions to many of
these, but, in most cases, definitive answers can be

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.6.

derived only from testing and experimentation. Not the
least of these problems is how best to organize the central
reference activities to make maximum use of the advantages
which EDP support can provide. While CHIVE has suggested
a number of organizational configurations which appear
to permit meeting system objectives as CHIVE sees them,
primary responsibility for decisions in this area must
rest with OCR. As yet no organization configuration has
been hit upon which solves all system problems. OCR
must judge CHIVE proposals in the light of its own
experience and with a view toward providing continuity for
major OCR functions. Hopefully, OCR will limit the
number of organizational options by adopting a "pre-
CHIVE" configuration which does not disrupt major func-
tions and yet permits experimentation without inhibiting
the development of the sound elements of the design and
without drastic and repeated overhauling of major com-
ponent outlines.

Another critical problem area is refinement of the
indexing procedures which, are as yet, excessively com-
plex and difficult to apply (see Volume V). An
extensive training period is required to master the index-
ing technique, and even after this period the indexer

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.6.

on occasion finds himself in doubt about how various segments of the technique are to be applied to differing problems. Reduction of the indexing complexity will be a major Phase III activity.

Closely tied to indexing problems is the matter of selectivity both of documents and of information within documents. It is clear at this point that available manpower cannot cope with the depth of indexing and the range of document categories anticipated in early stages of the design. The requirements of the various registers, however, are such that in an integrated system no simple exclusions by document category can be made without adverse effects on one or another of OCR's requrements. No major attack on this problem as yet has been undertaken, but it must be carefully studied during Phase III in close cooperation with the OCR divisions.

Still another major problem area is how the initial CHIVE component is to interact with the the OCR divisions as yet untouched by CHIVE. This involves not only such problems as duplication of indexing with geographic areas not yet included within CHIVE, or access to unconverted files, but the involvement of CHIVE in such scheduled outputs of OCR as the IPI and ███████████ publications.

25X1B

SUMMARY DESCRIPTION
Testing Concepts
3.1.3.6.

SECRET

System accessibility to uncleared requestors must also be examined during the Phase III activity. This matter will require careful study prior to any major investment in information file building. Safeguards to unauthorized disclosure of information must be constructed without placing limitations on the system's ability to provide service to customers who are not SI-cleared or who require information below the SI level. The capacity of the system to go as high as SI/T/KH or higher in answer to a query, but not to require the customer to accept data at this classification level whether he needs it or not, is a basic implementation objective.

SUMMARY DESCRIPTION
Testing Conecpts
3.1.3.6.

SECRET

Chapter 3.2.

IMPLEMENTATION TASKS

3.2.1. BASIC SCHEDULE

Given below is a timetable for passing major

implementation milestones during the Phase III effort:

3.2.1.1. Major Tasks (See chart following p. 35)

Date   Complete

Month 1   Phase II work and Phase III plan review completed
            EDP equipment ordered (IBM System/360 Model 60)
            Secure specifications of operating System/360
            (software)

            Begin

            Enlarged CSS on board
            ████████ Test Branch nucleus formed
            Begin building indexer aids
            Begin specification of information and summary
            files
            Begin definition of input/output procedures
            Program design underway
            Review of indexing procedures begins

Month 2   Complete

            Document Image System selected and ordered
            Analysis of OCR/CHIVE indexing experiment
            completed
            Specifications for page reader completed; page
            reader selected and ordered

            Begin

            Study of Map Library and Graphics Register
            inputs begins

                                    IMPLEMENTATION TASKS
                                    Basic Schedule
                                    3.2.1.1.

25X1A (left margin)

SECRET

Month 3   <u>Complete</u>

Forms design completed

<u>Begin</u>

Begin outlining dissemination procedures

Month 4   <u>Complete</u>

25X1A

███████ Test Branch T/O identified
EDP equipment (IBM 360/Mod 30) on site
Review of indexing procedures completed
Program design completed

<u>Begin</u>

Development programming begins

Month 5   <u>Begin</u>

Program unit testing begins

Month 6   <u>Complete</u>

Support file design completed

Month 7   <u>Complete</u>

Detailed design procedures for Document Image
System completed

Month 8   <u>Complete</u>

Document site selected and prepared

Month 9   <u>Complete</u>

Document Image System installed
Job sheets written for ██████   25X1A
Support files ready
Indexing procedures established
System test plan completed

<u>Begin</u>

System test plan development started

IMPLEMENTATION TASKS
Basic Schedule
3.2.1.1.

-33-

SECRET

25X1A

Month 10   <u>Complete</u>

Dissemination procedures completed
███ brought to full operational strength

<u>Begin</u>

Indexer training begins

Month 11   <u>Complete</u>

Initial information files fully defined
Basic operating System/360 available (software)
Page reader delivered and checked out

<u>Begin</u>

Start experimental indexing

Month 12   <u>Complete</u>

360/Mod 60 equipment installed

Month 13   <u>Complete</u>

System operating procedures completed

Month 14   <u>Complete</u>

System test development completed

Month 15   <u>Completed</u>

Indexing tests completed
Program unit testing completed
Graphics and Map Library inputs ready

<u>Begin</u>

Definition of interaction with OCR (projects,
inherited files)

Month 16   <u>Complete</u>

Secure full operating System/360 (software)

IMPLEMENTATION TASKS
Basic Schedule
3.2.1.1.

-34-

Begin

Final program checkout begins
Document dissemination started
Graphic and map inputs begin

Month 18   Begin

Detail procedures for customer access to ▊▊▊   25X1A
Branch
Customer training begins

Month 19   Begin

Operational testing begins

3.2.1.2.   Manning

The suggested manning requirements to accomplish
Phase III tasks are shown on the charts included below.

Manning requirements for an operational ▊▊▊ com-   25X1A
ponent would be substantially those of the ▊▊▊ Test   25X1A
Branch. In addition, four people (GS-4/6) will be required
to operate the Document Delivery System. Only the man-
power required for the design and implementation activities
is given here. Support manpower requirements, such as
key punch and computer operator manning, are summarized
in the Appendix 2A to Volume II of this report.

It must be recognized that manning levels, as assigned
to various tasks during initial planning, must be some-
what arbitrary and are subject to change as problem
areas are explored or as more realistic estimates of the

IMPLEMENTATION TASKS
Basic Schedule
3.2.1.2.

Figure 3-1 PERT Chart for CHIVE SYSTEM Development

MONTHS

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)

System Test Design

System Test Plan

System Test Development

Design of Support Files

System Operating Procedures

Analysis of Indexing Experiment

Approval of Phase III Plan

Support File Building

Selection and Procurement of the Document Delivery & System

Detailed Design Procedures for Document Delivery System Operations

Indexing Procedures Established

Experimental Indexing

System Test

Forms Design

Document Delivery System Site Preparation

Indexer Training

Specifications of Page Reader

Installation of Document Delivery System

Secure Specifications of Operating System /360

Program Design

Secure Basic Operating System/360

Program Unit Testing

Program Unit Test on an IBM System/360 on Site

Installation of Page Reader

Secure Full Operating System/360

IBM System/360 Procurement

Installation of IBM System/360 Model 60

Figure 3-2

TASK SCHEDULE

Task Key

1. Program Design
2. Development Programming
3. EDP System Test
4. Support Programming
5. Indexing Development
6. Data Gathering
7. System Procedures
8. File Building
9. Exper. Indexing and Training
10. System Integration
11. Operational Management
12. Design Management

Task schedule chart content:

- 1. Program Design: OCS-2; IBM-1 [3]
- 2. Development Programming: OCS-7; IBM-1 [8]
- 3. EDP System Test: OCS-7; IBM-3 [10]
- 4. Support Programming: OCS-1; IBM-2; CSS-1 [4] ... OCS-1; IBM-2; CSS-2 [5] ... OCS-1; CSS-3 [4]
- 5. Indexing Development: OCS-2; CSS-2; IBM-1; CTB-2 [7]
- 6. Data Gathering: IBM-1; CSS-1 [2]
- 7. System Procedures: OCS-1; IBM-1 CSS-1; CTB-4 [7] ... OCS-1; IBM-2 CSS-1; CTB-3 [7]
- 8. File Building: CTB-3 [3]
- 9. Exper. Indexing and Training: OCS-2; IBM-1; CSS-4; CTB-24 [31]
- 10. System Integration: OCS-1; IBM-1 [2] ... OCS-2; IBM-3 [5]
- 11. Operational Management: CTB-1 [1]
- 12. Design Management: OCS-1 [1]

Months: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18

-37-

Approved For Release 2000/05/08 : CIA-RDP78-03952A000100030001-9

Next 2 Page(s) In Document Exempt

amount of manpower required by certain tasks are reached. It has been assumed that skilled OCR personnel can be borrowed by CHIVE for short periods to assist in such activities as file building. As a result, manning requirements, as listed for some major tasks, may prove to be conservative.

3.2.2. PERSONNEL TASKS

In the following discussion it is assumed that Phase III task initiation and completion will adhere to the implementation time-table (or schedule of milestones) listed in the preceding section of this report. Tasks tend to break out into two closely related paths: 1) management tasks - those related to T/O problems, training, clearances, site location and preparation, procurement, dissemination, and relocation of people and their files; 2) substantive tasks - those involving preparation of indexer aids, refinement of index procedures, specification of CHIVE information and summary file formats, EAM file conversion, and data selection.

Management tasks relating to larger aspects of system implementation - equipment procurement, site preparation, data gathering, and OCR liaison and support

IMPLEMENTATION TASKS
Personnel Tasks
3.2.2.

should be the responsibility of the enlarged CSS working with

25X1A OCS and the ████ Test Branch management.

3.2.2.1. Substantive Personnel Tasks

The first implementation step could be to organize the
enlarged CSS. At the same time, the senior personnel who had
taken part in the Phase II Indexing Experiment, together with
a number of support personnel, and at least one EDP systems
analyst to advise on EDP possibilities, could be organized

25X1A into the nucleus of the future ████ Test Branch. The

25X1A personnel destined for the ████ Test Branch could be
organized on a task basis to work with CSS and the Design
Staff on the problems of preparing or improving the
basic indexer aids (organization dictionary, location
dictionary, etc.) needed for indexing in an operational
component.

Once the work on these basic tasks is well underway,
the indexer group--those responsible,for example, for
ISC revision and creation of new area rules--would devote
their main attention to working with OCS and CSS on
establishing and simplifying the indexing procedures,
setting up suggested indexing levels by document series,

25X1A and describing indexing levels for ████ data.
Presumably these tasks will be experimental and will be

IMPLEMENTATION TASKS
Personnel Tasks
3.2.2.1.

under revision and modification throughout the Phase III

activity. Experimental indexing should begin as soon

as the basic tools are constructed, the revised indexing

procedures defined, and interim EDP support made available.

25X1A ████████Branch, CSS, and OCS design personnel should

coordinate work on other major system procedures tasks

such as forms design, specification of summary and infor-

mation file input/output formats, maintenance and query

routines, customer access problems, and training.

3.2.2.2. Management Tasks

25X1A Management tasks for the outset of the ████████ Branch

activity would involve, first of all, the setting of

individual task schedules and the monitoring of

progress. Continual liaison with CSS and OCR must be

maintained on both substantive and administrative

matters. In the latter area the early identification

of OCR personnel to be detailed to file building is

of prime importance, and it is assumed that persons

with special skills or file knowledge may be borrowed

from time to time from OCR to aid in the creation of

indexer aids or summary files.

Once the initial increment of knowledgeable

personnel is found and engaged in system tasks, ████████ 25X1A

IMPLEMENTATION TASKS
Personnel Tasks
3.2.2.2.

Branch management can proceed to attack the problem of
the numbers and skills of the people to be involved
in the indexing activity, identify when their partici-
pation is to begin, arrange for clearances, and begin
preparation of draft job sheets describing the duties
to be performed and skills required for the personnel
25X1A of an operational ████ component. The proposed T/O
would be set during this period. Other matters requiring
detailed definition are the security procedures for the
operational testing phase, the establishment of dissem-
ination patterns, and the coordination of initial system
activities with the OCR registers, particularly in areas
where CHIVE activities and those of an OCR division
overlap.

Management tasks devolving on the CSS will include
advising on selection, procurement, site preparation,
and manpower problems relating to the Document Delivery
System, study of the problems of co-location of the
OCR document files, securing OCR support for data
collection and other design activities, and coordination
with OCS on EAM and EDP support requirements relating
25X1A to ████ activities.

IMPLEMENTATION TASKS
Personnel Tasks
3.2.2.2.

25X1A

3.2.2.3. ███ Organization

The manning of the CHIVE Support Staff has been

described above. Manpower for the ███ Test Branch   25X1A

can only be tentatively suggested at present since a

final determination will depend upon whether the test

component is established as an independent unit or is

attached, for administrative and logistical support,

to one of the extant OCR divisions. In either event,

25X1A

clericals would be required for header data indexing,

document control, filing and routine administrative

tasks. Specification of the exact manpower distribution

can be made only after final analysis of the results

of the OCR/CHIVE Indexing Experiment (see Volume V).

Since the primary aim of the ███ Branch is to   25X1A

launch a parallel program which, after a period of

experimentation, could become the first operational

IMPLEMENTATION TASKS
Personnel Tasks
3.2.2.3.

CHIVE increment, test indexing should being at the earliest feasible date. It is assumed that the full

25X1A complement of ███ members would be on board by month 10 so that ample time may be allowed for training.

3.2.2.4. Operator and User Training

Operator training, in its early stages, would repeat the training pattern established for the Phase II Indexing Experiment: three weeks of training on the revised ISC book, an OTR ███████████████ course

25X1A

if required, and an intensive training period in CHIVE indexing concepts and techniques. Rotational ISC training of OCR ████ analysts selected for later

25X1A

participation could take place well in advance of

25X1A their assignment to ████ At the conclusion of this training they could return to their home components while awaiting assignment to the ████ component.

25X1A

Training in a larger, practical sense would continue throughout the early experimental indexing stages

25X1A of ████ Branch activities. Training in the CHIVE query language would be started as soon as indexing personnel were well grounded in indexing procedures.

                              IMPLEMENTATION TASKS
                              Personnel Tasks
                              3.2.2.4.

A full scale program of communication with potential
system users would be undertaken during the weeks just
prior to assuming operational responsibility. A
special series of briefings could be given potential
users, and a short brochure describing the ██████Branch,   25X1A
its data base, files, and functions, could be prepared
and distributed widely. In the final operational check-
out period, during the last months of testing, many of
the ultimate system users could be involved in testing
and evaluating the system through query of the reference
and information files.

3.2.3. LOGISTICS TASKS

Logistical problems involving the computer center
are the responsibility of the Operations Division and
Administrative Staff of OCS and are not detailed in
this report except as noted in the milestone chart
given in Section 3.2.1.1.

3.2.3.1. Space

As indicated above, logistical tasks related to

25X1A   the ██████Test Branch and the Document Delivery System
would be the responsibility of OCR. The present 2G24 site
of the OCR/CHIVE Index Experiment provides a logical

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.1.

point from which to begin ███ Branch operations.   25X1A

Additional secure work space must be located and

prepared by month 10, however, in order to house

25X1A   the full ██ complement and its files, and to provide

some requestor work space. If possible, final space

selection for this component should take into

consideration its potential to expand to a full Far

Eastern division and its proximity to a central document

store. Total space required by the initial operational

component will be approximately 2000 sq. feet - 1670 sq. ft.

for personnel and approximately 330 sq. ft. for files.*

Initially, the Document Delivery System should require

no more than 675 sq. ft. of space in a secure area.

Assuming co-location of the present Library Circulation

Branch (Documents) and the SR Document Library with the

Document Delivery System, a total secured area of about

6,825 sq. ft. will be required.

The problem of the accessibility of the various

OCR document stores during the development of an all-source

25X1A   system will arise as soon as ███ Branch begins operational

---

*File space is required for indexer aids, some reference
works, and documents in various stages of processing. Final
file space requirements will depend, in part, on the operational
status of the microimage system.

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.1.

life. In preparation for this event, centralization of
the largest stores of documents and EAM files in OCR
could be considered as a way to provide a common document
repository out of which both the developing CHIVE system
and the areas not yet affected by CHIVE could work. If
the main library area on the first floor could be
revamped for this purpose, the SR documemt files, the
library hard copy and aperture card files and the new
document delivery system could all be located together.
The area on corridor 1H1, where the Librarian's office
is located, could also be studied as a possible site
for this common file. This area is also adjacent to the

25X1B ███ file rooms. Placing the first floor "H" corridor area

behind the SI barrier would permit free movement of system
personnel through this file space, and also allow access
to the nearby FIB files. Relocation of the large files
would permit central access to all major files by operators
of the old and new systems. It would also prevent repeated
relocation of file segments as new geographic areas are
added to the CHIVE initial component. Maintenance
and retrieval from the central file store could be
placed under a single management.

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.1.

EAM files and equipment might also be located in
one central area, possibly under a single management.
There would thus be a single point of address for EAM
support for non-CHIVE files or from inherited files,
and a single address point for document support whether
by system users or operators.

Area background files, special project manual
files, and smaller working files of interest to
particular groups of information analysts or particular
geographic areas would be housed with the information
analysts as system implementation progresses. CSS has
a full inventory of ███ files and file sizes (see
Volume V). An exact determination of total space
requirements can easily be made as soon as the size
and configuration of the initial system is agreed upon.

25X1B

3.2.3.2. Site Preparation

The site of the Document Delivery System must be
located and prepared by month 8 so that equipment
can be delivered and checked out. If centralization of
OCR document stores is undertaken, considerable reconstruc-
tion and preparation will be required in the area chosen
to house this facility. The same is true for EAM files
and machinery. CSS has obtained a plan of OCR floor

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.2.

space, including secure areas, and an inventory of
equipment and shelf and cabinet files, and could advise
and assist in site planning and preparation.

3.2.3.3. Auxiliary Equipment and Supplies

New equipment required for initial implementation
includes the Document Delivery System (basic card files,
cameras, printers, etc.) itself, and the optical page
reader designed for system inputs. Specially equipped
typewriters to prepare transcript sheets for acceptance
by the page reading devices will probably not be required.
The IBM "Elite" type appears to be the most common Agency
typewriter font, and the basic font for the page reader
and printers should probably be selected on the basis of
the availability of this font. Final selection of a type
font, however, and a possible subsequent requirement for
additional typewriters, will both be contingent on the
specifications established for the optical recognition
equipment. Other auxiliary equipment might include two
or more remote input/query stations (keyboard or CRT)
or, possibly, a long-distance Xerography link.

Whenever possible, equipment now owned or rented by
OCR or OCS will be used to support the initial system.
Examples of this would include Xerox 914 machines now
stationed in OCR, and OCR-owned film processors and

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.3.

keypunch equipment which can be utilized to support the Document Delivery System. A number of these are cited in the Appendix 2.A. in Volume II. The complete inventory of OCR equipment completed during the past year will be used to determine whether needed equipment is available or must be obtained.

The 360/Mod 60 computer system and attendant hardware is an OCS requirement, with or without CHIVE, and may therefore be considered available Agency equipment. Initially, OCR will share this computer with other users. However, the CHIVE machine-processing load should probably be reviewed after the first two years of operation to determine the best ultimate course to follow in managing the computer element of the system.

A number of remote communication terminals (IBM 1050 keyboard/printer, 1015 CRT Inquiry Display Terminal, and 2250 large CRT with light gun) are included in the system/360 equipment package being considered by OCS for installation in the Headquarters Building Computer Center. These terminals will be available for experimental purposes, and it is planned that one will be used to link

25X1A the � with the Computer Center. While these terminals will provide only a limited capability, as initially

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.3.

programmed, they can permit remote display of index

records or remote examination of query response while a

search is going on. Initial experimentation may help

in determining whether or not a total system requirement

for such remote display devices is likely to develop.

In a total system remote inquiry/output devices could

be used to link the geographic divisions with the Computer

Center or to connect either of these areas with such

remote requestor locations such as OO, NPIC, or OCI

Operations Center. Both ends of the remote terminal

will be in a secure area; security of the cable will

be explored as soon as a selection from the various

available remote devices has been made.

3.2.3.4. Move Schedules

Preparation of move schedules, beyond those needed

for relocation of the expanded ▇▇▇ by Month 9 or 10, and   25X1A

those required to prepare space for the Document

Delivery System, must await a decision on when the

initial system is ready to assume operational

responsibility. During a period of parallel operation,

colocation of any OCR files with a CHIVE component would

25X1A   work a hardship on current OCR operations.

Colocation of ▇▇▇▇▇▇▇ — — plant and town

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.4.

folders, or large document collections has not been
included in the basic task schedule. However, movement
of some of these files to a centralized location has been
recommended earlier in this volume. Designation of the
files to be relocated, a final estimate of their floor-
space requirements, identification of their containers,
preparation of new containers as needed, labeling of
files and their placement in the work area is all
comprised in this task. CSS should play an advisory
role in this task so that procedures for CHIVE access
to relocated files may be defined well in advance of
25X1A   operational activity by the ▮▮▮▮

Delivery of the Document Delivery System has been
scheduled tentatively for month 8. Moving of
files and clearing of floorspace to accommodate this
system and its personnel should be completed by late
October. Movement of files destined to be housed with
the Document Delivery System need not wait upon
installation and checkout of that system, but could be
scheduled for relocation as soon as the determined
25X1A   site was cleared of other occupants and secured for

IMPLEMENTATION TASKS
Logistics Tasks
3.2.3.4.

SECRET

3.2.4. DATA SPECIFICATION TASKS

This group of tasks involves the selection of the documents and document series to be indexed in the initial system, the identification and preparation of information and support files, the conversion of OCR EAM files required by the system, and establishment of selection criteria for the initial system indexers.

3.2.4.1. <u>Document</u> <u>Selection</u>

The document base and volume for ███████████████ 25X1A were determined prior to the Phase II CHIVE Indexing Experiment, and a description of this is given in Volume V of this report. Of the document categories known to be destined for inclusion in the initial system, those received through established Agency distribution channels; i.e., the basic USIB reporting, have been called "repository" documents. These will automatically receive header indexing. A number of document categories which are of interest to only certain information analysts ████████████████ STATSPEC

████████████████████████████ 25X1B

and journal clippings, etc.) have been called "non-repository" documents. These can be excluded from

IMPLEMENTATION TASKS
Data Specification
3.2.4.1.

SECRET

mandatory header indexing since these documents would

be selected and marked for indexing by the information

analysts. Only the "non-repository" documents selected

for content indexing and storage in the Document Delivery

System would require header indexing. Non-repository

documents which do not pass through regular dissemination

or are obtained by subscription, will pose a special

problem for the initial system. These documents will

have to be identified and copies directed to the ████   25X1A

The general flow of IR's, finished intelligence, cables,

STATSPEC FDD and ███ published materials should pose no

dissemination or procurement problem. However, a

number of document series (FDD scientific abstract

items, SI teletypes) which could eventually be received

in machine-readable form will have to be provided for.

The system personnel will have to (a) accept the idea

that these materials will not be indexed until they

can be machine prepared or (b) allow for additional

manpower and indexing time to input these documents,

or (c) make provision in Phase III to prepare these

inputs for system acceptance in machine language by the

end of Month 18. While the latter of these alternatives

is the most desirable, it will require special study

and the allocation of development manpower to make

IMPLEMENTATION TASKS
Data Specification
3.2.4.1.

these inputs ready by the start of ███ testing   25X1A
activities.

Disposition of TS documents and certain special
categories now controlled in SR will also require study
during the Phase III effort. A program for dealing with
these classes of documents will have to be worked out
before the start of Phase IV.

Even with information specialist marking of
documents for indexing, the document flow intended for
coding may prove too deep for available manpower. At
a minimum, present NODEX standards, and indexing levels
applied to various document series by the various
registers, must be examined to assist in determining
whether any major document series can be excluded from
indexing or, if indexed, to what level indexing must
be applied.

3.2.4.2. Data Selection

This is one of the most difficult of the tasks to
be undertaken during Phase III. A selectivity standard
based on current operational needs can be developed if
the information specialists are asked to only mark items
to be indexed in their documents and the indexing and
transcription of this material provided by a separate

IMPLEMENTATION TASKS
Data Specification
3.2.4.2.

group. Indeed, if this system proved satisfactory to the information retrieval personnel, it might never be necessary to specify a set of general system selection criteria for named object indexing except, perhaps, for some series. The danger in this approach is that the information specialist, with a free hand to mark items for someone else to process, may use very little discrimination in what he marks for indexer attention. A thorough review of processing criteria by geographic area for ███████████ and installational indexing would have to be made by the ████ and FIB management with selectivity criteria defined by the working personnel. FIB appears to have made a good start in this direction already, and ████ could probably formalize its processing requirements for ██████ as an initial attack on this problem.

Levels of data specification by document series, as currently applied in the various registers, would have to be summarized and compared with CHIVE minimum requirements and with the needs of the information registers. Some series could be handled at the header level only, still others would require only minimal named-object indexing. Others (diplomatic lists, plant

IMPLEMENTATION TASKS
Data Specification
3.2.4.2.

brochures, finished ████████ reports) could be given   25X1B

header indexing but could be retained in hard copy for

the topic or area to which they pertained.

3.2.4.3. Map Library and Graphics Register Inputs

During the Phase II activity, informal agreements

were worked out between OCS, ORR/Map Library, and OCR/

Graphics Register to provide for the development of

compatible index standards which would enable either

the CHIVE analyst, or the graphics or map analysts, to

search the central computer system for index records to

maps, photos and films. A major Phase III task will

be the coordination of index development, the establish-

ment of index levels, and arriving at index compatibility

with Map Library and GR so that computer inputs can be

ready for acceptance by October 1966. Development

Division personnel will be assigned to this task and

persons familiar with the GR and Map Library systems

and requirements will, presumably, be loaned to the

development activity by the map and graphics components.

It is not planned that the map and graphics document

files be made a part of a central document delivery system.

The physical proximity of these files to the central system

is, however, desirable. GR is already located in an area

IMPLEMENTATION TASKS
Data Specification
3.2.4.3.

adjacent to the first floor location which has been suggested as a possible site of a central document store. Conceivably, the small Map Library Headquarters collection could be located nearby.

3.2.4.4. Vocabulary Control Files

The construction of indexer tools and vocabulary control files would be among the very first implementation tasks. These are discussed below.

3.2.4.4.1. Area Dictionary

The area rule problems as applied during the CHIVE Indexing Experiment will be examined. Intelligence Subject Code area rules would be modified and a special area indexing procedure generated especially for CHIVE's indexing problems. A Master Location Dictionary for

25X1A ▮▮▮▮ based on the ISC dictionary and NIS Gazetteer would be developed.

3.2.4.4.2. Intelligence Subject Code

Suggested revisions to the ISC book would be collected and incorporated into a CHIVE version of the ISC. In addition, the ISC would be expanded to include those subject/codes peculiar to Comint materials.

IMPLEMENTATION TASKS
Data Specification
3.2.4.4.2.

3.2.4.4.3. Occupation and Function Code

Special occupation and organization function listings for personality and installation indexing will be developed. The ISC in its present form would not be used for this purpose.

3.2.4.4.4. Organization Identifier File

25X1A

A ███████ Organization Identifier File will be developed consolidating the organizational information available in SR, FIB, and ███ cross-referenced and 25X1B non-redundant, and reformatted for computer. This would probably be the most sizeable of the file building tasks and could be expected to extend through the first six months of the Phase III effort. Specialists

25X1A

on ████████ organizations would be borrowed from appropriate OCR divisions to advise during this activity. One of these people would probably be identified as the organization dictionary editor of the future ██████████ 25X1A Branch. At least two subtasks would be involved in this effort: (a) establishing a liaison procedure with SR,

25X1B

FIB and ████ to pick up newly identified ██████ organizations 25X1A and to record changes in established ones, and (b) analysis of the political organization problem peculiar to ████████ 25X1B category files to determine whether these files are

IMPLEMENTATION TASKS
Data Specification
3.2.4.4.4.

suitable for incorporation in the basic organization
dictionary or whether they are more efficiently maintained
in manual form.

25X1B   3.2.4.4.5. ███████████████████ File

    The possibility of reformatting the ███████ listing 25X1B
for use both as an index to ███████████████ and as 25X1B
a basic reference aid must be studied. The entire card
deck for each ████████ entry could be rearranged so that 25X1B
all elements of information coded to the ██████ could 25X1B
be listed after the name entry.

3.2.4.4.6. Header Data Dictionary

    Creation of a more detailed Header Data Dictionary
is required to establish standardized control terminology
for recording and displaying data contained in the
bibliographic or header portion of documents.

3.2.4.5. File Conversion

    Present planning calls only for initial conversion
of those EAM-based OCR files which are required by CHIVE
to provide vocabulary control aids for the system. These
are listed in the preceding section. Conversion of the
EAM card base for the Intellofax, SR, FIB, or ████████ 25X1B
document or ███████████ presents a number of severe 25X1B

IMPLEMENTATION TASKS
Data Specification
3.2.4.5.

problems and will require careful examination of the
most feasible conversion routes for the most active
files. This matter is being studied and some further
discussion of the problem and its approaches is given
in Volume V.

3.2.4.6. Information Files

EDP-based summary information files would be
created by the information specialists assigned to the
25X1A   ▮▮▮▮ Test Branch. These files would be begun on the
basis of an established requirement for the maintenance
25X1B   of such files in ▮▮▮ FIB, or any other OCR component.
25X1A   An all-source ▮▮▮ would face at once the problems of
25X1B   all-source ▮▮▮▮▮ and installation summary files.
These would include organizational and target installation
listings, some of which may be identical with or derived
from, identifier lists. Construction of non-redundant,
25X1B   summary ▮▮▮▮▮▮▮ and installation summaries
could be begun as soon as experimental indexing was
started. Construction of these files would make use,
wherever possible, of work already done in ▮▮ and FIB.   25X1B
Other summary files, such as ▮▮▮▮▮▮ (a contribu- 25X1B
tion only), leader appearances, travel listings, may
also be required. In addition the initial system must

~ IMPLEMENTATION TASKS
Data Specification
3.2.4.5.

be prepared to simulate or specify procedures for the
preparation of finished intelligence reports and meeting
production deadlines for the products cited above.

The system should also experiment with the preparation
of summary information files from unsynthesized data
in the index records, and should attempt to gain experience
in preparation of these outputs. Timeliness, both of
input and output, will be a significant factor in judging
the success of the unsynthesized file approach in filling
requirements for such closely scheduled system products
as leader appearance listings.

Information files will be constructed at the lowest
possible classification level. Here, the problem of
all-source inputs will be a significant factor, since
the files must be ready to serve a wide range of system
customers who will want products at the all-source level
and below.

3.2.5. PROCEDURES TASKS

The tasks described below are related to the problems
of how the system is to operate once it enters an
experimental indexing mode and thereafter. A number of
activities touched on in preceding sections, experimental

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.

use of remote input/inquiry devices, for example, might
also be considered procedures tasks.

3.2.5.1.   Indexing/Transcription Procedures

An indexing guide was developed for the OCR/CHIVE
Indexing Experiment (see Appendix 5.C). For an operational
system the manual needs considerable expansion to cover:

- Indexing procedures unique to ███ as an area.   25X1A

- Special indexing rules which should be developed
  out of the experience gained in the experiment.

- Document selection procedures unique to ███ as   25X1A
  an area.

Header data and content data transcription sheets
have also been developed, but experimental use indicates
some need for revision. A single header data transcription
sheet was developed for all sources for use in the experiment.
It has been recommended that during Phase III several forms be
developed to accommodate various series and that/the use
of continuous roll forms be investigated as an aid to
expedite typing. Both the header data form and the
content data form which is used to retype the content
data from the indexer's handwritten transcription form
will have to be developed with the requirements for
character reader input in mind.

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.1.

3.2.5.2. Other Forms Design Requirements

3.2.5.2.1. Summary File Forms

25X1B   ████████████ and installation summary files may require special input/output forms. The content requirements of these files can be established by the appropriate OCR divisions. However, considerable study of formatting problems may be required.

3.2.5.2.2. Transaction Processing

A general form on which to record information about processing activity will also be required. A version of the input transcription form can be passed through the page reader for computer action; however, additional data about the query transaction must be captured. Possibly some variation of the data collection forms used in the OCR/CHIVE Indexing Experiment can be employed initially and revised as new requirements come to light.

3.2.5.2.3. Document Requests

Another form will be required for communication with the Document Delivery System when large numbers of documents or files are requested and it would be burdensome to transmit document numbers by telephone.

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.2.3.

3.2.5.2.4. Error Correction

An error correction form may also be required so that items in index records rejected as incorrect by the computer can be corrected by the indexer with a minimum of reindexing or retyping.

3.2.5.2.5. Index Instruction

A form may be needed through which the information analyst can specify his input or maintenance instructions to the indexing staff. In many instances, marking the documents themselves may not be feasible. This area of analyst/indexer communications is, as yet, unexplored territory, and detailed procedures specification in this area will become a major system requirement.

3.2.5.3. Operator/User Communications

No novel operator/user communication techniques are planned for the initial system. Communication with system customers will take place by mail, by personal customer visit, by black and gray phone communication, and through use of the Headquarters Building tube systems. Agency mail and courier service will be used

25X1A   between ███ and remote locations.

The major task in the communications area will be

25X1A   to define and simplify ███ contact with customers who

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.3.

are not cleared at the all-source level. A special work

area will have to be established for customers who

cannot visit the Branch itself, and ███ analysts will

have to meet with customers in this area to discuss

requests, make files available, and explain query products.

25X1A

Establishing feedback procedures to record and

describe customer reactions to the developing system

will be another essential system task. Feedback can be

captured verbally, by CHIVE analyst comment on his

requestor's reactions, and also by followup forms attached

to distributed query products. Circulation of a

questionnaire to known system customers at the conclusion

of the initial operational test period will also be

considered.

3.2.5.4. Maintenance/Query Procedures

An extensive set of file maintenance and query

capabilities has been designed for the initial EDP

system. Appendix 7.A. describes the query language

by which the information analyst will request the

performance of these functions by the EDP system. The

optical page reader will accept the user coding sheets

for maintenance and query as well as for index record

input. Experimental usage of the query language will

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.4.

be required to determine whether improvements can be made in its ease of use and capabilities. Possible changes might involve altering the syntax of the language, the coding forms, or, most importantly, the functions performed by the language.

The structuring of new files, new output reports, and new input formats will be carried out initially by the programming staff by way of the assembly/compiler languages of the system/360. As additional operating experience is gained, higher level, problem-oriented languages will be developed for the structuring of files, inputs, and reports.

3.2.5.5. Management Reporting Procedures

Throughout the implementation phase it is assumed that both periodic and ad hoc reports to OCR and OCS management will be made through CSS, ▮▮ and Development   25X1A
Division personnel. Written reports will be produced as various milestones are passed or as management approval for design decisions is required. ▮▮ will   25X1A
report routine management information on work accomplished, personnel actions, etc., in the manner of a line OCR component. As soon as experimental indexing has begun, logical products of this activity, e.g., volume figures,

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.5.

indexer rates, key punch times, and transaction speeds can be collected. The initial component will have the capacity to produce manually any of the data now reported routinely to OCR management by its operating components, and to generate automatically information on machine activity of interest to OCS management and to design personnel.

3.2.5.6. Procedures Documentation

Input, maintenance, and output procedures for the CHIVE document index files will be summarized and described throughout the initial implementation period. Indexing manuals will be created and kept under revision. In addition, a query guide, which will also include access and maintenance routines for summary files, vocabulary control files, and other identifier listings, will have to be drawn up. Capabilities of and access to external OCR files of interest to the CHIVE system will also be described, as will procedures for interaction or feed-in to on-going OCR programs which may be influenced by CHIVE developments.

IMPLEMENTATION TASKS
Procedures Tasks
3.2.5.6.

SECRET

3.2.6. PROGRAMMING TASKS

3.2.6.1. Program Design

The CHIVE System Design has specified an overall
system flow, system loads (input, processing, and
output), system priorities, system queues, etc. This
information will be used in Phase III for program design.

The program designers will specify the logical
program units that must be written. These program
specifications will delineate all of the relevant
parameters associated with each program unit. These
parameters include;

- input data formats, devices and volumes;

- processing requirements for file building,
  maintenance and query;

- output devices, volumes, and formats;

- storage allocations; and

- interface between program units.

3.2.6.2. Development Programming

This task involves the translation of the CHIVE
Program Design into an operational program. The first
step in accomplishing this task is to prepare flow
charts of all EDP functions. These flow charts will be
reviewed by project management to insure that the

IMPLEMENTATION TASKS
Programming Tasks
3.2.6.2.

-71-

system design has been properly interpreted. The next step is coding into an appropriate language which can be read by a computer processor. Testing will then be performed on program units which will later be assimilated into program modules for further testing. While a program unit performs limited functions such as sorting, merging, and report structuring, these units will be combined into program modules (such as output processing, file searching, and file maintenance) for further testing.

3.2.6.3. System Testing

During the program development task a set of objective criteria will be established to be used as a final test of the programming system. This test will demonstrate to management the capabilities of the CHIVE Operational program.

The complete program system must finally be tested for proper interaction among the program modules and to insure that the program units and modules properly interface with the operating system. Input conditions are established and data used to simulate operational conditions. This testing will uncover flaws in the program design, flow charting, and coding efforts. These errors will be corrected under this task.

IMPLEMENTATION TASKS
Programming Tasks
3.2.6.3.

The final phase of testing will consist of using the previously developed set of criteria for system testing in order to secure management approval of the systems operation.

3.2.6.4. Support Programming

The development of an operational EDP system requires a variety of computer programs other than the operational program itself. Support programming on CHIVE has been carried on for several months in support of the document indexing experiment. This effort will be continued and will expand to include the programs necessary to convert existing files into formats that can be efficiently processed by a computer as well as the development of support files such as authority lists. The capabilities of existing programs will be exploited before special purpose programs are written since efficiency of program operation is not a significant factor in this task.

3.2.6.5. IBM System 360 Software

Major decisions have yet to be reached on the degree to which the more sophisticated IBM programming packages can be exploited in the development of the CHIVE EDP system. The target dates for availability of these packages are such that severe delays could be encountered in providing the initial capability if full reliance on

IMPLEMENTATION TASKS
Programming Tasks
3.2.6.5.

-73-

IBM software is planned. While reliance on these programs is advisable because of their power and the need for Agency programming standards, the need to get on with the job may require that some initial work be undertaken without these programs being available.

IMPLEMENTATION TASKS
Programming Tasks
3.2.6.5.

SECRET